Physiotherapy rehabilitation for osteoporotic vertebral fracture—a randomised controlled trial and economic evaluation (PROVE trial)

K.L. Barker*[1,2], M. Newman[2], N. Stallard[3], J. Leal[4], C.M.Lowe[2], M.K. Javaid[1], A. Noufaily[3], T. Hughes[2] D. Smith[1], V. Gandhi[1], C. Cooper[1] and S.E. Lamb[1] on behalf of the PROVE trial group

* karen.barker@ouh.nhs.uk

[1]Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences (NDORMS), Botnar Research Centre, University of Oxford, Oxford OX3 7LD, UK

[2]Physiotherapy Research Unit, Physiotherapy Department, Nuffield Orthopaedic Centre, Oxford University Hospitals NHS Foundation Trust, Windmill Road, Oxford OX3 7LD, UK

[3]Division of Health Sciences, Warwick Medical School, University of Warwick, Coventry, UK

[4]Health Economics Research Centre, Nuffield Department of Population Health, University of Oxford, Oxford, UK

Abstract

Summary

The trial compared three physiotherapy approaches: manual or exercise therapy compared with a single session of physiotherapy education (SSPT) for people with osteoporotic vertebral fracture(s). At 1 year, there were no statistically significant differences between the groups meaning there is inadequate evidence to support manual or exercise therapy.

Introduction

To evaluate the clinical and cost-effectiveness of different physiotherapy approaches for people with osteoporotic vertebral fracture(s) (OVF).

Methods

Prospective, multicentre, adaptive, three-arm randomised controlled trial. Six hundred fifteen adults with back pain, osteoporosis, and at least 1 OVF participated. Interventions: 7 individual physiotherapy sessions over 12 weeks focused on either manual therapy or home exercise compared with a single session of physiotherapy education (SSPT). The co-primary outcomes were quality of life and back muscle endurance measured by the QUALEFFO-41 and timed loaded standing (TLS) test at 12 months.

Results

At 12 months, there were no statistically significant differences between groups. Mean QUALEFFO-41: − 1.3 (exercise), − 0.15 (manual), and − 1.2 (SSPT), a mean difference of − 0.2 (95% CI, − 3.2 to 1.6) for exercise and 1.3 (95% CI, − 1.8 to 2.9) for manual therapy. Mean TLS: 9.8 s (exercise), 13.6 s (manual), and 4.2 s (SSPT), a mean increase of 5.8 s (95% CI, − 4.8 to 20.5) for exercise and 9.7 s (95% CI, 0.1 to 24.9) for manual therapy. Exercise provided more quality-adjusted life years than SSPT but was more expensive. At 4 months, significant changes above SSPT occurred in endurance and balance in manual therapy, and in endurance for those ≤ 70 years, in balance, mobility, and walking in exercise.

Conclusions

Adherence was problematic. Benefits at 4 months did not persist and at 12 months, we found no significant differences between treatments. There is inadequate evidence a short physiotherapy intervention of either manual therapy or home exercise provides long-term benefits, but arguably short-term benefits are valuable.

Trial registration ISRCTN 49117867.

Introduction

Osteoporosis is a major public health problem that affects an estimated 3.2 million people in the United Kingdom (UK).[1] It's clinical importance lies in its association with bone fractures and their complications.[2] Vertebral fractures are the most common osteoporotic fracture and are thought to affect at least 20% of the older population in the UK; thus around 1 in 5 people aged 50 years or more will have one or more osteoporotic vertebral fractures (OVF), with the incidence and prevalence increasing rapidly with age.[2, 3]

OVFs result in increased mortality and significant morbidity, even if initially clinically asymptomatic.[3-5] They are associated with back pain, fatigue, impaired mobility, depression, restricted activity and social participation and marked reductions in quality of life (QoL) that can persist for at least 18 months post-fracture. [4-6] Vertebral fractures cause spinal deformity and alter spinal biomechanics, increasing the risk of subsequent vertebral fractures, hyperkyphosis, falls, non-vertebral fractures and restrictive lung disease.[4,5,7] Osteoporosis and fragility fractures also present a substantial economic burden; with the cost of care in the UK estimated to rise from £4.4 billion pounds in 2010 to approximately £5.5 billion pounds per year by 2025.[1]

Physiotherapy is frequently recommended for patients with symptomatic OVFs and can include a variety of exercise (strength, balance, postural exercises etc) and 'hands-on' manual therapies such as joint mobilisations and massage.[1,7,9-11] In conjunction physiotherapists often provide substantial education e.g.; about osteoporosis, fall prevention strategies etc.[12] However, evidence of effectiveness is lacking, including which type of physiotherapy might be most beneficial and the

costs of treatment.[9] The aim of this trial was to assess the clinical and cost-effectiveness of three different physiotherapy approaches: manual therapy and exercise therapy for people with symptomatic OVF(s) compared with a single session of individually tailored education from an experienced physiotherapist.

Methods

This was a multicentre, prospective, three-arm randomised controlled trial (RCT) with an adaptive design and blinded outcome assessment. It was a pragmatic trial, designed to measure the effectiveness of physiotherapy interventions in standard conditions. It took place in 21 National Health Service hospital outpatient physiotherapy centres across England. The trial was approved by the UK South Central Research Ethics Committee (REC: 12/SC/0411), and site-specific approvals were obtained. It was performed according to the published protocol, with no changes after the trial began[13]. The study and intervention protocols were designed in accordance with the CONSORT and TIDieR guidelines.

Participants

Potential participants were approached by staff at osteoporosis clinic visits or by mail, using electronic medical records from relevant radiology or primary care clinics to identify those with OVF. They were given an invitation letter and information pack and those interested were offered a research appointment. Men and women aged 18 or over were eligible if they had a diagnosis of osteoporosis confirmed by radiograph or dual energy X-ray absorptiometry (DEXA) scan (T score -2.5 SD below young adult mean at the lowest lumbar level), at least one previous OVF (confirmed on existing radiography, magnetic resonance imaging or DEXA scan) and back pain. Individuals at

different times post-fracture, with different numbers of fractures and fracture locations were eligible. They also had to be able to walk at least 10 metres, be able to exercise and participate in physiotherapy safely and, if female, to be post-menopausal. People with severe unstable cardiovascular or pulmonary disease, significant psychiatric or neurological conditions, bone loss secondary to other disease or medication, those who had undergone vertebroplasty, facet joint injection or physiotherapy in the previous 12 weeks were ineligible.

Randomisation and masking

Research clinicians confirmed participant eligibility, obtained written consent and conducted baseline assessment prior to randomisation to ensure allocation concealment. The central telephone service at Warwick Clinical Trials Unit randomised participants between a program of exercise therapy, manual therapy or a single session of physiotherapy (SSPT) education and advice using a variable block randomisation schedule stratified by centre to control for confounding factors at local recruitment sites. Research staff did not deliver treatments, and all staff conducting assessments and managing data were blinded to allocation group at all points. It was not possible to mask participants or physiotherapists providing interventions, but participants were asked not to disclose their allocated group to assessors.

Outcome Assessments

Clinical outcome assessment occurred at baseline, 4 and 12 months. The primary endpoint was 12 months. The co-primary outcomes were: 1) QUALEFFO-41 : a disease specific health related QoL self-report questionnaire with 5 domains (pain, physical function, social function, general health perception, mental function), total score range 0 – 100, lower scores indicate higher QoL,[14] and 2)

the Timed Loaded Standing test (TLS) which records the time (in seconds) a person can stand with arms extended, shoulders flexed to 90^0 holding a 1kg weight.[15] Holding time and test end relate to back muscle endurance.[15,16] Secondary outcomes included thoracic kyphosis angle measured with a flexicurve ruler,[17] standing balance evaluated using the Functional Reach test (FR), physical function and walking capacity assessed by the Short Physical Performance Battery (SPPB), a 6-minute walk test (6MW) and Physical Activity Scale for the Elderly (PASE), and falls incidence recorded with a monthly event calendar.[18-21] Participants completed the EQ5D-5L and logged healthcare use in a standardised trial diary. Attendance and adverse events were recorded on standardised trial forms.

Interventions

Participants in manual therapy and exercise therapy were offered a 1-h assessment plus 6 individual outpatient physiotherapy sessions spread over 12 weeks with a specialist musculo-skeletal physiotherapist with at least 4-year post-graduate musculoskeletal experience. Interventions were individually tailored for each participant. Participants allocated SSPT were offered a single 1-h assessment, education, and advice session with a physiotherapist. See online Appendix 1 and published protocol for further detail.[13] The comparator was single one-hour assessment, education and advice session with a physiotherapist. SSPT comprised general education about osteoporosis, vertebral fractures and strategies to reduce falls. It covered lifestyle choices to promote bone health including general information about diet, regular weight-bearing exercise and physical activity. Information was consistent with osteoporosis clinical guidelines and information from the Royal Osteoporosis Society (ROS, UK).[2,11,23] Further detail is in the published protocol[13].

Manual therapy

Manual therapy included low velocity central posterior-anterior spinal mobilisations through the thoracic and/or lumbar spine, with vertebral level, grade (from grade 2 to 4) and number of repetitions selected individually. It also included soft tissue mobilisation, postural education including taping, and a home programme of passive stretches that promoted thoracic extension for up to 15 min daily[13].

Exercise therapy

Exercise therapy consisted of a multi-component, progressed programme of balance, strength training concentrating on back extensor and postural muscles and functional weight-bearing exercise (walking, step-ups, etc.)[13]. Based on assessment, the physiotherapist selected exercises which were sufficiently challenging and could be performed effectively, safely, and comfortably from standardised, graded sets. Strength training intensity was set and monitored using the 0–10 rating of perceived exertion (RPE) scale at a participant-perceived moderate to somewhat hard level of effort (RPE level 3, 4)[22]. Balance exercises were progressed in a standard way, e.g. upper limb support to no upper limb support and eyes open to eyes closed. A pedometer was used to assess walking capacity, structure, and progress the walking programme[3,22]. Participants were asked to continue the exercise programme at home between clinic sessions aiming to include short sessions of exercise within daily life, aiming to achieve 45 min per day, 3 to 5 times a week depending on ability. Strategies were employed to promote adherence[13].

Single session of physiotherapy

This consisted of general education about osteoporosis, vertebral fractures, and strategies to reduce falls. It covered lifestyle choices to promote bone health including general information about diet,

regular weight-bearing exercise, and physical activity. Information was consistent with osteoporosis clinical guidelines and information from the Royal Osteoporosis Society (ROS, UK)[2,11,23]. The participants in the manual and exercise groups received general advice only.

Physiotherapists received 4 h of training covering theoretical and practical application of the interventions. Therapists detailed the content of all treatment sessions in a treatment log and their usual clinical records. The research team reviewed treatment logs and visited each site to conduct fidelity checks, monitoring implementation against study protocols using a proforma and ensuring each therapist received a minimum of one quality control visit per treatment type that they gave. Further visits occurred based on the outcome of the initial visit.

Sample size

No formally established minimal clinically important difference (MCID) exists for QUALEFFO-41 or TLS tests. The sample size was calculated to detect a standardised effect of 0.4 in the QUALEFFO-41; at 80% power with an alpha of 0.05, which would require 180 to 200 participants in each group. We calculated between 540-600 participants were required, allowing for 10% drop-out.

Statistical Analysis

This was an adaptive trial, in which either active intervention arm could be dropped or the trial halted at an interim analysis. The interim analysis was guided by pre-specified rules using the estimated mean change in total QUALEFFO-41 from baseline to 4 months using data from 75 participants per arm. If change from baseline in an intervention arm was no more than 0.5 points greater than SSPT, this arm would be dropped. Under this rule, both arms might be dropped, and the trial terminated. A change in one arm, more than 2 points higher than the other, would result in dropping the worst performing arm. Considering this, the specified sample size gave 94% power if

the better of the two intervention arms had a true standardised effect of 0.4. After accrual and follow-up of the first 75 participants, the rules were met for continuing as a three-arm trial.

There was a pre-specified statistical analysis plan. Intention to treat analyses was performed, and participants were analysed in accordance with the group to which they were randomised. Baseline characteristics for participants in each arm were summarised by proportions for each level binary or categorical variables and means and standard deviations for continuous variables. We pre-specified that we would not impute missing data if the proportion of participants with missing data was less than 10% within each arm. The analyses of primary and secondary outcomes were adjusted for centre and baseline value of the outcome being considered. The analyses of primary outcomes were additionally adjusted to account for the adaptive design of the study and for the multiple comparisons arising from the two pairwise comparisons with SSPT using a closed testing procedure. We used an inverse normal combination function to combine individual and Dunnett-corrected p values obtained from treatment comparison adjusting for centre and baseline from the datasets from the two stages. No adjustment for multiplicity was made in analysing secondary endpoints. For the primary outcomes, unbiased estimates and confidence intervals correcting for the adaptive design and multiple testing were also obtained[24]. Results are presented as treatment effects with 95% confidence intervals and p values; statistical significance was defined as p = 0.05 (two-sided 5% significance level).

Pre-planned sub-group analyses investigated the interaction between age (70 versus > 70 years), sex, baseline fracture status (2 versus > 2), and treatment allocation on primary outcomes adjusted for centre and baseline value. A pre-specified complier-average causal effect (CACE) analysis using two-stage least squares estimation was completed to determine if attendance affected primary

outcomes using the number of sessions attended as an endogenous variable and adjusting for baseline and recruitment centre as exogenous variables. All analyses used R version 3.4.1 (www.R-project. org) except the CACE analysis, which used STATASE 15.0 (www.stata.com).

The cost-effectiveness analysis used EQ5D-5L and healthcare diary data collected from participants. The perspective was that of the NHS and personal social services (price year 2015/2016) and the outcome was incremental cost per quality-adjusted life year (QALY). The time horizon of the analysis was 1 year and costs and QALYs were not discounted. Missing resource use and EQ5D data were dealt with using multiple imputation where data were missing at random using predictive mean matching with five nearest neighbours. The difference in costs and QALYs was estimated with regression analysis, using a system of seemingly unrelated regressions. The probability of each intervention being the most cost-effective was estimated at a threshold of £20,000 per QALY gained[25]. Further details on the cost-effectiveness analysis are available in the online appendix, reported using the CHEERS checklist.

Adverse events

These were reported to the trial team and investigated further. They were classified through discussions with the local Principal Investigators, the Trial lead and one of the co-applicants Dr. Muhammad K Javaid, a Consultant in Metabolic Bone Medicine and reviewed by the trial Data Monitoring and Efficacy Committee (DMEC).

Patient and Public Involvement

Throughout, we were supported by the Royal Osteoporosis Society a UK charity, and a representative from this organisation was a co-applicant. Members of our local osteoporosis group and former patients with osteoporosis were also consulted when developing and refining trial interventions and materials. In a pilot phase at trial outset, a nested qualitative study explored the acceptability of the trial to participants. Another patient representative was a voting member of the trial steering committee and provided guidance.

Results

From September 2013 to September 2016, 1213 potentially eligible people were identified of which 615 were enrolled. In total, 63/ 615 (10%) participants withdrew and a further 17/615 (3%) were non-contactable at 4 months, rising to 23/615 (4%) at 12 months (Figure 1). The most common reason for withdrawal was death or serious other illness: exercise (n=9), manual therapy (n=11), SSPT (2). Dissatisfaction with allocated arm was common in SSPT (n=7) but not in manual (n=1) or exercise therapy (n=0). Final data was collected in September 2017.

Participants were aged from 42 to 97 years, mean age 72 (SD 9.1) years; 531/ 613 (86%) were women, 82/613 (13%) were men. The mean lumbar spine DEXA T-score was -2.7 (SD 1.3), the average number of OVF was 2.5 (SD 1.9), most participants were hyperkyphotic and had substantially limited mobility and endurance (Table 1). The groups were well-matched at baseline. Loss to follow up was highest in the exercise arm but less than 10% in all arms. Primary data at 12 months was obtained for 529/615 (86%) participants for QUALEFFO-41 and 458/615 (75%) for TLS. Those without QUALEFFO-41 had slightly higher pain, those without TLS were more likely to have back pain currently, worse balance and kyphosis.

Intervention compliance

One participant withdrew on randomisation to SSPT, the remaining 195/196 (99%) attended. Exercise participants attended a mean 4.3 (SD 2.7) sessions; 143/216 (66%) partially complied (minimum 4 sessions) and 82/216 (38%) fully complied (7 sessions). Manual therapy participants attended a mean 5.03 (SD 2.6) sessions; 155/203 (76%) partially complied and 99/203 (49%) fully complied. Fidelity checks showed no evidence of contamination (that participants crossed arms) and that intervention content was well delivered. Common barriers to attendance were other health problems, caring commitments, and transport difficulties. Healthcare diaries showed that 50/196 (25%) SSPT participants sourced some type of additional physiotherapy (NHS or private) outside of the trial.

Adverse Effects

A total of 85 adverse events were reported but there were no serious adverse events associated with the trial, according to the pre-specified criteria. In exercise, 24/216 (11%) participants reported 26 events including 5 falls and 6 fragility fractures; in manual therapy, 34/203 (17%) participants reported 37 events with 6 falls and 9 fragility fractures, in SSPT 22/196 (11%) reported 22 events including 4 falls and 8 fragility fractures. Other events included the following: cardiovascular problems (stroke, tachycardia, etc.), respiratory and urinary tract infections, minor musculoskeletal pain (shoulder, back), and other diverse conditions.

Clinical Outcomes

At 12 months there was no statistically significant effect observed from either intervention over SSPT on either primary outcome. Sub-group and CACE analyses showed no significant differences. Within groups total QUALEFFO-41 scores improved slightly relative to baseline and TLS endurance increased by a mean 9.8s, or 20% (exercise), a mean 13.6s, or 28% (manual therapy) and 4.2s or 8% (SSPT). Whilst these changes were clinically significant, there were no statistically significant findings between the groups. Improvements in thoracic kyphosis were largest in the manual and exercise therapy groups over 12-months (Table 2). Mean kyphosis reduction at 12 months was − 6.9° (exercise), − 4.6° (manual therapy), and − 2.7° (SSPT), changes that are recognised as clinically significant, but there were no statistically significant differences between the groups.

Although there was no statistically or clinically significant difference in outcomes at 12 months, there was some difference in the early pattern of response, measured at 4 months, with clinically and statistically significant improvements in the primary outcome of TLS endurance in manual therapy compared to SSPT and in those 70 years or younger in exercise therapy. The CACE analysis suggested improvements due to manual therapy increased with the number of sessions attended. In manual therapy compared with SSPT, TLS increased by a mean of 1.95 s (95% CI 0.45 to 3.44), p = 0.01 for each session attended. In-home exercise, TLS increased by a mean of 1.52 (95% CI − 0.23 to 3.27), p = 0.09 for each session attended. At this point, there were also significant improvements in balance (FR) in both exercise and manual therapy above SSPT and significant improvements in functional mobility (SPPB and 6 MW) above SSPT in exercise therapy (Tables 2 and 3).

Cost-effectiveness

Exercise therapy was more effective (0.002 QALYs, 95% CI − 0.020 to 0.025) but more costly than SSPT (£206, 95% CI − £228 to £641) whereas manual therapy was less effective (− 0.014, 95% CI − 0.036 to 0.009) and more costly than SSPT (£244, 95% CI − £195 to £683) (Table 4 and Table A.3

Online appendix). Using a threshold of £20,000 per QALY gained, the most cost-effective option was SSPT with a probability of 69% (Table 4). Sub-group analyses indicated exercise therapy to be cost-effective for those aged 70 years or younger (£12,310 per QALY gained) but not for those over 70 (52% probability). Full data on resource use, costs, EQ5D utility, and cost-effectiveness analysis will be published in the National Institute for Health Research health and technology assessment journals library.

Discussion

Overall, at 12 months, this trial found no statistically significant differences between the three treatments and SSPT to be the most cost-effective option. All of the treatments demonstrated some degree of benefit that would be consistent with clinically important changes, but overall, there was no significant difference between them.

The reduction in thoracic kyphosis in the exercise and manual therapy arms at 12 months was of a level judged clinically important in other trials.[10,26] Thoracic kyphosis usually progresses with age and hyperkyphosis is associated with excess morbidity and mortality, so even small reductions are valuable if sustained.[8,10,26,27] Significant differences in back muscle endurance as measured by timed loaded stand were seen early after treatment but by 12 months, there was no significant difference between the groups, although modest increases persisted. Back muscle fatigue and difficulty carrying out tasks in standing due to fatigue are key problems for people with OVF and chronic back pain that affect QoL[15,28]. Back muscle weakness is also correlated negatively with kyphosis angle, associated with hyperkyphosis and an increased risk of fractures[29–31]. Most participants had poor endurance on trial entry,[15,16] and deterioration or no improvement would be expected in this population over 12 months[29].

The sub-group analysis highlighted improvements in endurance were experienced primarily by those 70 years or younger. Ageing is associated with increased osteoporosis severity, number of comorbidities and frailty, and neuro-muscular changes such as sarcopenia. These factors may compound problems due to spinal pathology for older people with osteoporosis.[28] Younger participants may have had more capacity to improve and/or found home exercise and treatment attendance easier.

At 4 months, balance improved significantly in both intervention arms above SSPT. A 1-cm increase in FR in older adults is thought clinically relevant;[32] mean changes greater than 2 cm occurred in both arms, the larger in exercise therapy potentially due to combined balance and strength training. In addition, significant, clinically relevant gains in functional mobility and walking capacity were seen in exercise therapy[33]. By 12 months, improvements in balance and mobility in intervention groups reduced in size, gains were seen in SSPT, and differences were no longer significant.

Comparison with other studies

Over the trial there were no significant differences in QoL between groups as measured by total QUALEFFO-41 or EQ5D-5L. However, the improvement we saw within exercise therapy was of similar magnitude to significant findings in a trial without an active comparator.[12] Evstigneeva et al. (2016) reported larger benefits immediately after a 12-month (104 session) intervention, but here QoL deteriorated in the no-intervention control, emphasising effects and no longer follow-up occurs. [35] Five other RCTs report no significant change in total QUALEFFO-41 between groups after treatment, although one shows significant change in subscales and one in an accompanying generic QoL measure. [10, 36-39] Papaioannou et al. (2003) studying a frail cohort similar in age and disease

severity to ours, found QoL improved after 6 months of home-based physiotherapy and exercise and a further 2 RCTs of exercise interventions in women with and without OVFs demonstrate improvements in QoL post-treatment. [40-42] These studies use other osteoporosis-specific QoL instruments. [40-42] Although QUALEFFO-41 is the most common measure, whether it is the best instrument for evaluating change due to rehabilitation is unclear.[43]

Increased back strength or endurance immediately after physiotherapy has been observed in other RCTs in this population.[39, 41, 44, 45] Improved balance immediately after treatment is also reported. [12, 35] Bergland et al. (2011) found improvements in FR of a comparable magnitude after a 12 week (24 session) exercise class intervention.[12] Improvements have been associated with reduced fear of falling,[46] but as in our trial not with significant reductions in falls at 1 year. [37, 46] The two other studies that conduct longer follow-up found, like us, when supervised intervention ceased differences between groups diminished and at 12 months changes were not significant.[12, 44]

Consistent with our trial, three previous RCTs of exercise therapy for people with OVFs have demonstrated significant improvements in functional mobility post-treatment. [12, 35, 46] In contrast, Papaioannou et al. (2003) found no significant improvement.[40] Whether exercise dose was sufficient is uncertain, adherence to 3 days/ week of home exercise was 62% (23/37 women) at 6 months, decreasing to 46% (17/37) at 12 months.[40] In our trial, mobility gains reduced and were no longer significant at 12 months. Unlike us, Bergland et al. (2011) found gains persisted at 12 months.[12] Here, exercise dose appears higher, 75% (34/45 women) attended at least 79% (19) of 24 classes, and the group format may have provided additional benefits.[12]

Strengths and limitations of this study

This trial is the largest RCT to assess physiotherapy for people with OVFs to date and the first to provide information about cost effectiveness. Most participants had severe osteoporosis; characterised by multiple OVFs, hyperkyphosis, back pain, poor back muscle endurance and limited exercise capacity. The trial was well conducted and pragmatic, designed to test interventions deliverable in routine practice. Contrary to most previous studies it evaluated longer term effects of interventions.

However, our intervention was potentially too low intensity to sustain improvements. Determining the intervention dose was challenging, models of treatment delivery in RCTs in similar populations vary widely, e.g. from individual sessions plus daily home-based self-directed exercise, to 104 class sessions over 12 months[34, 40]. Although 7 sessions across 12 weeks seem limited when compared with prolonged exercise class interventions, previous trials with individual sessions and home programmes reported benefits, and individual sessions are required for manual treatments.[10, 12, 34–41, 43, 44] Our interventions resembled these trials and tested a common model of service delivery, i.e. outpatient physiotherapy rehabilitation plus a self-directed home programme that was feasible to deliver within a publicly funded healthcare system.

An issue specific to the cost-effectiveness analysis was the missing EQ5D and patient diary data. Complete EQ5D-5L and cost data was available for only 242/613 (40%) of participants. However, the sensitivity analysis explored the impact of different imputation models and SSPT remained the most cost-effective intervention at 12 months.

Our comparator was education by a physiotherapist and 25% (50/196) participants in SSPT sourced additional physiotherapy of some kind outside of the trial. Our results show that unlike trials with no treatment comparators that report deterioration, small gains were seen in SSPT in most outcomes, against the expected disease course and suggestive of static success. These gains may represent a placebo response, educational benefits, or the effect of physiotherapy treatment sourced outside of the trial but reduced between-group difference. Furthermore, most participants did not receive the planned intervention intensity, potentially resulting in smaller treatment effects in both intervention groups. This was highlighted by the CACE analysis which suggested those who attended more sessions experienced greater benefits. Other studies of both class and individual interventions have also reported problems with compliance[10, 38, 43, 44, 46]. Direct comparisons are complex as compliance is defined and measured in varied ways across studies. For example, Bautmans et al. (2010) report 15/29 (52%) participants comply, attending 50% (9/18) physiotherapy sessions over 12 weeks, Gold et al. (2004) noted 94 participants attended an average of 58% (70/120) exercise classes, or see Papaioannou et al. (2003) (discussed above)[10, 38, 43]. Multiple factors affect treatment adherence in older people with chronic diseases such as osteoporosis; problems with transport, caring commitments, and ill-health were common barriers to attendance reported by participants in our trial, but other personal and programme characteristics might be influential[47]. Overall, these factors potentially resulted in smaller effects in the intervention groups and magnified effects in SSPT, exposing the trial to the risk that intervention effects were lost or underestimated.

Further research

Further research is important into factors that assist compliance, alongside investigations that consider whether age and intervention dose affect outcomes. Studies are also needed to evaluate

the responsiveness of the QUALEFFO-41 to rehabilitation interventions compared to other osteoporosis QoL instruments and to establish a formal MCID.

Conclusions and implications

The key messages are that at the main endpoint at one year the trial did not find significant differences between the three treatments in terms of effectiveness. However, several measures showed benefit at 4-months and some benefits persisted at 12-months. Whilst there is not adequate evidence to support manual or exercise therapy for long-term benefit it is arguable that short-term benefits are valuable and that education with a physiotherapist confers benefit.

Contributions of Authors

Karen L Barker (Professor of Physiotherapy) Chief Investigator, led funding application, trial conception and design, development of interventions, supervision, writing and reviewing report.

Meredith Newman (Research Fellow) Trial lead, trial design, development of interventions, supervision, writing and reviewing report.

Nigel Stallard (Statistician) Designing and conducting statistical analysis, writing and reviewing report.

Jose Leal (Health Economist) Designing and conducting economic analysis, supervision, writing and reviewing report.

Catherine Minns-Lowe (Research Associate, Qualitative): Qualitative study design and conduct, writing and reviewing report.

Muhammad K Javaid (Associate Professor, Consultant in metabolic bone disease) Trial supervision, trial design, assessment of adverse events, reviewing report.

Angela Noufaily (Statistician) Conducting statistical analysis, writing and reviewing report.

Anish Adhikari (Health Economist) Economic analysis, writing and reviewing report

Tamsin Hughes (Research Physiotherapist) Trial intervention and assessments, writing and reviewing report

David J Smith (Trial co-ordinator) Trial management, writing and reviewing report.

Varsha Gandhi (Trial co-ordinator): Trial management, quality assurance of interventions, trial management, writing and reviewing report.

Cyrus Cooper (Professor of Rheumatology) Trial conception and design, trial supervision, writing and reviewing report.

Sarah E Lamb (Professor of Rehabilitation; Director of Clinical Trials Unit): Trial conception and design, trial supervision, writing and reviewing report.

The Trial Steering Committee: Professor Maria Stokes (chairperson), Professor Sarah Lamb, Dr Victoria Allgar, Dr Allison Rushton, Dr Jane Simmonds, Dr Fiona Cramp, Mrs Jane Aldridge (patient representative)

Data Monitoring Committee: Professor David Torgerson (chairperson), Professor Susan Todd, Professor Helen Dawes.

Site Principal Investigators and Research Clinicians

Charlotte Heywood, Phillip Clarke (Addenbrookes Hospital), Heather Shilliday (Countess of Chester Hospital), Dr Carol McCrum, Kate Weatherly (Eastbourne General Hospital, Conquest Hospital), Nicky Wilson (Kings College Hospital), Ira Pande, Marie-Josephe Pradere, Rachel Nicole (Queens Medical Centre), Tamsin Hughes (Nuffield Orthopaedic Centre), Kim Brown, Anna Thornhill, Emma McLoughlin (Queen Alexandra Hospital), Sarah Moore, Liz Jacobs, Stephanie Travers-Griffin (Heavitree Hospital), David Rogers, Gareth Stephens (Royal Orthopaedic Hospital), Rebecca Hull, Joanne Richardson, Michael Dawson (Royal Surrey County Hospital), Julie Russell, Sophie Park (Royal National Hospital for Rheumatic Diseases), Jessica Shipley, Jacqueline Harrison (Northern General Hospital, Royal Hallamshire Hospital), Steven Young Min (Queen Alexandra Hospital, St Mary's Hospital), Linda Knott, Nicola Donlin, John Clare (Torbay District General Hospital), Hubert van Griensven, Dr Billy Fashanu, Steven Young, Joanne Calver (Southend Hospital), Zoe Paskins Lucy Huckfield (Haywood Hospital), David Cumming, Sajeed Kaleel, Maggie Dawson, Sheeba Suresh, Mary Knott (The Ipswich Hospital), Jonathan Marks, Emily Watson (The Royal Bournemouth Hospital, Christchurch Hospital), Sophia Mavrommatis (University College Hospital), Nicola Clague, Ross Dixey (Leicester General Hospital, Leicester Royal Infirmary), Chris Mercer, Isobel Amey (Worthing Hospital).

Funding Support

Funding information This research was funded by the National Institute for Health Research (NIHR Health Technology Assessment Programme—HTA 10.99.01). NIHR funded the study. It developed a commissioning brief but had no role in the study design, data acquisition, analysis, or manuscript

preparation. The chief investigator (KLB) assumes overall responsibility for the data in the manuscript. All named investigators had full access to the trial dataset. The views expressed are those of the authors and are not an official view of the institutions or funders.

Data Sharing Statement

All data requests should be submitted to the corresponding author for consideration. Access to available anonymised data may be granted following review.

References

1. Svedbom A, Hernlund E, Ivergård M, Compston J, Cooper C, Stenmark J, et al. Osteoporosis in the European Union: a compendium of country-specific reports. *Arch Osteoporos* 2013; 8:137.doi.org/10.1007/s11657-013-0137-0

2. Compston J, Cooper A, Cooper C et al. UK clinical guideline for the prevention and treatment of osteoporosis. *Arch Osteoporos* 2017;12:43. doi.10.1007/s11657-017-0324-5.

3. Ballane G, Cauley JA, Lackey MM, El-Hajj Fulcihan G. Worldwide prevalence and incidence of osteoporotic vertebral fractures. *Osteoporos Int* 2017; 28: 1531-1542. doi:10.1007/s00198-017-3909-3

4. Kendler DL, Bauer DC, Davison KS et al. Vertebral fractures: clinical importance and management. *Am J Med* 2016;129:221 e1-10.

5. Sanchez-Riera L, Wilson N. Fragility fractures & their impact on older people. *Best Pract Res Clin Rheumatol* 2017;31:169-91.

6. Svedblom A, Borgström F, Hernlund E, Ström O, Alekna V, Bianchi ML, et al. Quality of life for up to 18 months after low-energy hip, vertebral, and distal forearm fractures—results from the ICUROS. *Osteoporos Int* 2018; 29:557-566 doi.org/10.1007/s00198-017-4317-4

7. Broy SB. The vertebral fracture cascade: Etiology and clinical implications. *J Clin Densitom* 2016;19:29-34.

8. van der Jagt-Willems HC, de Groot MH, van Campen JPCM, Lamoth CJC, Lems WF. Associations between vertebral fractures, increased thoracic kyphosis, a flexed posture and falls in older adults: a prospective cohort study. *BMC Geriatrics* 2015; 15. doi:10.1186/s12877-015-0018-z.

9. Giangregorio LM, MacIntyre NJ, Thabane L, Skidmore CJ, Papaioannou A. Exercise for improving outcomes after osteoporotic vertebral fracture. *Cochrane Database Sys Rev* 2013, Issue 1. Art. No.: CD008618. doi: 10.1002/14651858.CD008618.pub2.

10. Bautmans I, Van Arken J, Van Mackelenberg M, Mets T. Rehabilitation using manual mobilization for thoracic kyphosis in elderly postmenopausal patients with osteoporosis. *J Rehabil Med* 2010;42:129-35. doi: 10.2340/16501977-0486

11. Kanis JA, Cooper C, Rizzoli R, Reginster JY on behalf of the Scientific Advisory Board of the European Society for Clinical and Economic Aspects of Osteoporosis (ESCEO) and the Committees of Scientific Advisors and National Societies of the International Osteoporosis Foundation (IOF). European guidance for the diagnosis and management of osteoporosis in postmenopausal women. *Osteoporos Int* 2019; 30:3–44. doi.org/10.1007/s00198-018-4704-5

12. Bergland A, Thorsen H, Karesen R. Effect of exercise on mobility, balance, and health related quality of life in osteoporotic women with a history of vertebral fracture: a randomized, controlled trial. *Osteoporos Int* 2011;22(6):1863-71. doi:10.1007/s00198-010-1435-7

13. Barker KL, Javaid MK, Newman M, Minns Lowe C, Stallard N, Campbell H, et al. Physiotherapy Rehabilitation for Osteoporotic Vertebral Fracture (PROVE): study protocol for a randomised controlled trial. *Trials* 2014;15:22. doi.org/10.1186/1745-6215-15-22

14. International Osteoporosis Foundation (IOF). Quality of Life Questionnaires QUALEFFO-41. https://www.iofbonehealth.org/quality-of-life-questionnaires-qualeffo-41. Accessed September 5, 2018.

15. Shipp KM, Purse JL, Gold DT, Pieper CF, Sloane R, Schenkman M, et al. Timed loaded standing: a measure of combined trunk and arm endurance suitable for people with vertebral osteoporosis. *Osteoporos Int* 2000;11(11):914-22.

16. Newman M, Newman R, Hughes T, Vadher K, Barker KL. Is the timed loaded standing test a valid measure of back muscle endurance in people with vertebral osteoporosis? *Osteoporos Int* 2018;29:893-905. doi:10.1007/s00198-017-4358-8.

17. Greendale GA, Nili NS, Huang MH, Seeger L, Karlamangla AS. The reliability and validity of three non-radiological measures of thoracic kyphosis and their relations to the standing radiological Cobb angle. *Osteoporos Int* 2011;22:1897-905.

18. Duncan PW, Studenski S, Chandler J, Prescott B. Functional reach: predictive validity in a sample of elderly male veterans. *J Gerontol* 1992;47(3):M93-8.

19. Guralnik JM, Simonsick EM, Ferrucci L et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. *J Gerontol* 1994; 49(2):M85-94.

20. Steffen TM, Hacker TA, Mollinger L. Age- and gender-related test performance in community-dwelling elderly people: Six-Minute Walk Test, Berg Balance Scale, Timed Up & Go Test, and gait speeds. *Physical Therapy* 2002;82:128-37.

21. Washburn RA, Smith KW, Jette AM, Janney CA. The Physical Activity Scale for the Elderly (PASE): development and evaluation. *J Clin Epidemiol* 1993;46:153-62.

22. ACSM. Exercise Prescription for people with osteoporosis. In: Petit MA, Hughes JM, Warpeha JM, editors. ACSM resource manual for guidelines for exercise testing and prescription. 6th ed 2010.

23. National Osteoporosis Society (NOS). About osteoporosis: Your bone strength. https://nos.org.uk/about-osteoporosis/your-bone-strength/. Accessed September 5, 2018.

24. Friede T, Parsons N, Stallard N et al. Designing a seamless phase II/III clinical trial using early outcomes for treatment selection: an application in multiple sclerosis. *Stat Med* 2011;30:1528-40.

25. Dunn G, Maracy M, Tomenson B. Estimating treatment effects from randomized clinical trials with noncompliance and loss to follow-up: the role of instrumental variable methods. *Stat Methods Med Res* 2005;14:369-95.

26. Royal Osteoporosis Society (ROS). Strong, steady and straight: An expert consensus statement on physical activity and exercise for osteoporosis. https://theros.org.uk/forms/documents/strong-steady-and-straight/. Accessed April 1, 2019.

27. Katzman WB, Vittinghoff E, Lin F et al. Targeted spine strengthening exercise and posture training program to reduce hyperkyphosis in older adults: results from the study of hyperkyphosis, exercise, and function (SHEAF) randomized controlled trial. *Osteoporos Int* 2017;28:2831-41.

28. Kado DM, Miller-Martinez D, Li-Yung L et al. Hyperkyphosis, kyphosis progression, and risk of non-spine fractures in older community dwelling women: The study of osteoporotic fractures (SOF). *J Bone Miner Res* 2014; 29: 2210-2216.

29. da Silva RA, Vieira ER, Cabrera M et al. Back muscle fatigue of younger and older adults with and without chronic low back pain using two protocols: A case-control study. *J Electromyogr Kines* 2015;25:928-36.

30. Ball JM, Cagle P, Johnson BE, Lucasey C, Lukert BP. Spinal extension exercises prevent natural progression of kyphosis. *Osteoporos Int* 2009; 20: 481-489.

31. Sinaki M, Itoi E, Wahner HW et al. Stronger back muscles reduce the incidence of vertebral fractures: A prospective 10 year follow-up of postmenopausal women. *Bone* 2002;30:836-41.

32. Mika A, Unnithan VB, Mika P. Differences in thoracic kyphosis and in back muscle strength in women with bone loss due to osteoporosis. *Spine* 2005;30(2):241-6.

33. Idland G, Rydwik E, Smastuen MC, Bergland A. Predictors of mobility in community dwelling women aged 85 and older. *Disabil Rehabil* 2013;35:881-7.

34. Perera S, Mody SH, Woodman RC, Studenski SA. Meaningful change and responsiveness in common physical performance measures in older adults. *J Am Geriatr Soc* 2006;54:743-9.

35. Evstigneeva L, Lesnyak O, Bultink IE et al. Effect of twelve-month physical exercise program on patients with osteoporotic vertebral fractures: a randomized, controlled trial. *Osteoporos Int* 2016;27:2515-24. doi:10.1007/s00198-016-3560-4

36. Liu-Ambrose TY, Khan KM, Eng JJ, Lord SR, Lentle B, McKay HA. Both resistance and agility training reduce back pain and improve health-related quality of life in older women with low bone mass. *Osteoporos Int* 2005;16:1321-9. doi 10.1007/s00198-005-1842-3

37. Grahn Kronhed A-C, Hallberg I, Ödkvist L, Möller M. Effect of training on health related quality of life, pain and falls in osteoporotic women. *Adv Physiother* 2009;11:154-65.

38. Carter ND, Khan KM, McKay HA et al. Community-based exercise program reduces risk factors for falls in 65- to 75-year-old women with osteoporosis: randomized controlled trial. *CMAJ* 2002;167(9):997-1004.

39. Bennell KL, Matthews B, Greig A et al. Effects of an exercise and manual therapy program on physical impairments, function and quality-of-life in people with osteoporotic vertebral fracture: a randomised, single-blind controlled pilot trial. *BMC Musculoskel Disord* 2010;11:36. http://www.biomedcentral.com/1471-2474/11/36.

40. Papaioannou A, Adachi JD, Winegard K et al. Efficacy of home-based exercise for improving quality of life among elderly women with symptomatic osteoporosis-related vertebral fractures. *Osteoporos Int* 2003;14: 677-82. doi: 10.1007/s00198-003-1423-2

41. Hongo M, Itoi E, Sinaki M et al. Effect of low intensity back exercise on quality of life and back extensor strength in patients with osteoporosis: a randomized controlled trial. *Osteoporos Int* 2007;18:1389-95. doi: 10.1007/s00198-007-0398-9

42. Madureira MM, Bonfá E, Takayama L. Pereira RMR. A 12-month randomized controlled trial of balance training in elderly women with osteoporosis: Improvement of quality of life. *Maturitas* 2010; 66: 206–211. doi:10.1016/j.maturitas.2010.03.009

43.Madureira MM, Ciconelli, RM, Pereira RMR. Quality of life measurements in patients with osteoporosis and fractures. *Clinics* 2012; 67: 1315-1320.

44.Gold DT, Shipp KM, Pieper CF, Duncan PW, Martinez S, Lyles KW. Group treatment improves trunk strength and psychological status in older women with vertebral fractures: Results of a randomized, clinical trial. *J Am Geriatr Soc* 2004;52:1471-8.

45. Bergström I, Bergström K, Kronhed A-CG, Karlsson S, Brinck J. Back extensor training increases muscle strength in postmenopausal women with osteoporosis, kyphosis and vertebral fractures. *Adv Physiother* 2011;13:110-7. doi: 10.3109/14038196.2011.581696

46.Olsen CE, Bergland A. The effect of exercise and education on fear of falling in elderly women with osteoporosis and a history of vertebral fracture: results of a randomized controlled trial. *Osteoporos Int* 2014; 25:2017–2025. doi: 10.1007/s00198-014-2724-3

47.Kerschan-Schindl K, Uher E, Kainberger F, Kaider A, Ghanem AH, Preisinger E. Long term home exercise program: Effect in women at high risk of fracture. *Arch Phys Med Rehabil* 2000;81:319-23.

Tables

Table 1: Baseline Participant Characteristics

	Exercise Therapy	Manual Therapy	SSPT
No. (%) female participants	185 (85.6%) (n=216)	173 (85.6%) (n=202)	173 (88.7%) (n=195)
Age mean (SD), y.	72.2 (8.4) (n=216)	72.4 (9.3) (n=203)	71.9 (9.6) (n=196)
Height mean (SD), cm	156.4 (20.6) (n=216)	159.1 (8.7) (n=201)	155.6 (20.8) (n=195)
Weight mean (SD), kg	64.3 (16.4) (n=216)	64.1 (14.4) (n=202)	62.9 (10.9) (n=195)
DEXA lumbar T score, mean (SD)	-2.6 (1.5) (n=162)	-2.7 (1.2) (n=162)	-2.8 (1.3) (n=147)
No. spinal fractures, mean (SD)	2.7 (1.8) (n=196)	2.4 (1.8) (n=187)	2.5 (2.1) (n=169)
No. non-spinal fractures, mean (SD)	0.2 (0.4) (n=194)	0.2 (0.4) (n=184)	0.1 (0.3) (n=177)
No. (%) back pain last 2 weeks	209 (96.8%) (n=216)	194 (96.0%) (n=202)	188 (96.4%) (n=195)
No. (%) back pain today	157 (72.7%) (n=216)	133 (65.8%) (n=202)	128 (66.0%) (n=194)
NPRS back pain, mean SD (0 -10)	4.8 (2.2) (n=156)	4.5 (2.1) (n=133)	4.6 (2.1) (n=127)
No. falls in last year, mean (SD)	0.8 (2.3) (n=214)	0.7 (1.4) (n=201)	0.6 (1.0) (n=192)
No. fallers (%) in category 1[(a)]	1 (0.5%) (n=207)	1 (0.5%) (n=193)	0 (0.0%) (n=188)
category 2	19 (9.2%) (n=207)	14 (7.3%) (n=193)	8 (4.3%) (n=188)
category 3	187 (90.3%) (n=207)	178 (92.2%) (n=193)	180 (95.7%) (n=188)
QUALEFFO-41, mean (SD)	39.9 (16.0) (n=214)	37.1 (14.9) (n=200)	38.1 (15.9) (n=195)
TLS, mean (SD), s	48.6 (54.5) (n=210)	47.9 (51.7) (n=197)	52.4 (58.7) (n=193)
PASE, mean (SD)	109.7 (89.7) (n=191)	115.6 (81.3) (n=182)	106.9 (71.2) (n=165)
Thoracic kyphosis mean (SD), deg.	51.2 (34.8) (n=213)	48.6 (22.4) (n=199)	49.1 (40.6) (n=194)
SPPB, mean (SD)	8.3 (2.1) (n=200)	8.6 (2.1) (n=185)	8.4 (2.2) (n=184)
Functional reach mean (SD), cm	23.5 (8.8) (n=216)	23.4 (10.5) (n=201)	23.8 (9.8) (n=195)
6 MWT, mean (SD), m	295.4 (128.0) (n=216)	304.2 (135.3) (n=201)	313.4(129.4) (n=194)
Abbreviations: DEXA, dual energy X-ray absorptiometry, NPRS, numeric pain rating scale, TLS, timed loaded standing test, PASE, physical activity scale for the elderly, SPPB, short physical performance battery, 6MWT, 6-minute walk test, (a)Participant falls 1 = Frequently (≥ once a week), 2 = Occasionally (≤ once a month), 3 = Rarely (≤ once a year)			

Table 2: Primary and Secondary Outcomes for Exercise Therapy and Manual Therapy vs SSPT

Outcome	Time	Exercise Therapy			Manual Therapy			SSPT
		Mean (SD) change from baseline	Change relative to control adjusted for centre and baseline (95% CI)	P value	Mean (SD) change from baseline	Change relative to control adjusted for centre and baseline (95% CI)	P value	Mean (SD) change from baseline
Primary Outcomes								
QUALEFFO-41 total (0-100) *	4 mo.	-2.16 (8.40) (n=180)	-1.23 (-2.84,0.37)	p=0.13	-0.55 (8.17) (n=185)	0.27 (-1.32,1.86)	p=0.74	-0.87 (7.05) (n=173)
	12 mo.	-1.31 (9.97) (n=176)	-0.23 (-3.20, 1.59)[+]	p=1.00[+]	-0.15(10.81) (n=181)	1.35 (-1.76, 2.93)[+]	p=0.74[+]	-1.18(8.89) (n=172)
TLS s	4 mo.	7.42 (47.90) (n=162)	7.46 (-1.43,16.35)	p=0.10	11.80 (40.27) (n=171)	10.56 (2.22,18.90)	p=0.01	-0.58 (40.90) (n=161)
	12 mo.	9.8 (52.4) (n=148)	5.77 (-4.85,20.46)[+]	P=0.44[+]	13.6 (38.9) (n=153)	9.69 (0.09,24.86)[+]	P= 0.33[+]	4.2 (55.8) (n=157)
Secondary Outcomes								
QUALEFFO-41* Pain	4 mo.	-7.76 (18.51) (n=180)	-3.57 (-7.39, 0.25)	p=0.07	-5.58 (21.41) (n=182)	-2.54 (-6.56, 1.48)	p=0.27	-3.77 (19.13) (n=173)
	12 mo.	-9.07 (22.88) (n=176)	-2.80 (-7.07, 1.46)	p=0.19	-6.75 (22.98) (n=180)	-1.03 (-5.22, 3.17)	p=0.63	-6.25 (18.39) (n=172)
Physical function	4 mo.	-2.81 (9.45) (n=180)	-1.02 (-2.83, 0.80)	p=0.27	-0.89 (8.54) (n=185)	0.73 (-0.99, 2.45)	p=0.41	-1.66 (8.35) (n=173)
	12 mo.	-0.70 (11.41) (n=176)	0.18 (-2.04, 2.41)	p=0.87	1.10 (11.56) (n=181)	1.78 (-0.48, 4.04)	p=0.12	-0.90 (10.16) (n=172)
Social function	4 mo.	-3.22 (14.32) (n=180)	-2.71 (-5.46, 0.03)	p=0.05	-0.83 (15.75) (n=185)	-0.98 (-3.86, 1.90)	p=0.41	-0.10 (13.30) (n=172)
	12 mo.	-1.65 (16.09) (n=175)	-0.66 (-3.68, 2.37)	p=0.67	0.02 (17.26) (n=181)	-0.11 (-3.31, 3.08)	p=0.94	-0.45 (16.11) (n=172
General health perception	4 mo.	-1.99 (15.48) (n=180)	-1.07 (-4.01, 1.88)	p=0.48	0.43 (15.09) (n=184)	0.65 (-2.17, 3.48)	p=0.65	0.10 (13.45) (n=173)
	12 mo.	-0.17 (18.83) (n=175)	-0.06 (-3.61, 3.50)	p=0.97	1.00 (18.15) (n=180)	0.07 (-3.33, 3.47)	p=0.97	1.03 (16.09) (n=170)
Mental health	4 mo.	-0.23 (10.32) (n=180)	1.56 (-0.59, 3.71)	p=0.16	-0.84 (10.50) (n=184)	0.96 (-1.17, 3.09)	p=0.37	-2.06 (11.27) (n=173)
	12 mo.	1.45 (11.69) (n=174)	1.65 (-0.81, 4.10)	p=0.19	0.22 (13.14) (n=180)	0.18 (-2.35, 2.71)	p=0.89	-0.37 (12.56) (n=171)
Thoracic Kyphosis* deg.	4 mo.	-4.21 (40.63) (n=170)	0.82 (-3.27,4.90)	p=0.70	-2.58 (9.31) (n=173)	-0.71 (-3.96,2.55)	p=0.67	-3.26 (40.33) (n=167)
	12 mo.	-6.88 (31.66) (n=152)	-2.24 (-6.40,1.92)	p=0.29	-4.63 (22.28) (n=156)	-3.02 (-8.14,2.10)	p=0.25	-2.73 (49.58) (n=167)
PASE	4 mo.	10.26 (54.58) (n=133)	0.9 (-13.58,15.48)	p=0.90	9.38 (86.20) (n=136)	1.67 (-16.78,20.12)	p=0.86	8.02 (66.41 (n=125)
	12 mo.	9.72 (63.64) (n=116)	6.33 (-8.40,21.06)	p=0.40	4.22 (67.02) (n=127)	1.56 (-12.92,16.04)	P=0.83	6.75 (57.58) (n=122)
SPPB	4 mo.	0.81 (1.72) (n=152)	0.45 (0.11,0.79)	p=0.01	0.31 (1.51) (n=157)	-0.03 (-0.35,0.30)	p=0.88	0.39 (1.57) (n=150)
	12 mo.	0.66 (1.78) (n=134)	0.18 (-0.21,0.57)	p=0.36	0.15 (1.49) (n=141)	-0.19 (-0.54,0.16)	p=0.28	0.47 (1.75) (n=149)
Functional Reach cm	4 mo.	1.28 (9.19) (n=170)	2.43 (0.71,4.16)	p=0.01	1.53 (9.56) (n=177)	2.11 (0.57,3.64)	p=0.01	-0.80 (9.85) (n=168)
	12 mo.	0.26 (8.57) (n=152)	1.02 (-0.96,3.01)	p=0.31	1.18 (13.00) (n=155)	0.54 (-1.65,2.73)	p=0.63	-0.21 (12.72) (n=165)
No. Falls in last year [(a)]	4 mo.		-0.08 (-0.46,0.29)	p=0.66		-0.05 (-0.41,0.31)	p=0.78	
	12 mo.		-0.09 (-0.41,0.24)	p=0.60		-0.17 (-0.49,0.16)	p=0.32	

6MWT m	4 mo.	28.09 (102.5) (n=167)	26.09 (6.58,45.60)	p=0.01	16.73 (90.18) (n=174)	16.16 (-1.76,34.09)	p=0.08	-1.73 (84.71) (n=166)
	12 mo.	19.1 (103.77) (n=150)	6.83 (-13.08, 26.74)	p=0.50	20.19 (87.45) (n=157)	12.02 (-6.46,30.50)	p=0.20	7.16 (88.48) (n=164)
*A negative change corresponds to a better outcome, a) estimate reported is log-risk ratio [+]Adjusted for adaptive design and multiple testing								

Table 3: Sub-group analyses

	Exercise Therapy		Manual Therapy		No. participants (Exercise; Manual Therapy; SSPT)
	Change relative to SSPT adjusted for centre and baseline (95% CI, p-value)	P value for interaction	Change relative to SSPT adjusted for centre and baseline (95% CI, p-value)	P value for interaction	
QUALEFFO-41 at 12 months					
Female	-0.41 (-2.49, 1.66) p=0.695	0.357	0.87 (-1.27, 3.00) p=0.428	0.985	(154;155;154)
Male	3.61 (-2.09, 9.30) p=0.225		0.42 (-7.69, 8.54) p=0.919		(22;26;18)
Age ≤ 70	-0.00 (-3.22, 3.22) p=1	0.834	-0.36 (-3.92, 3.20) p=0.842	0.937	(68;66;68)
Age ≥7 0	-0.58 (-3.13, 1.97) p=0.656		0.83 (-1.73, 3.38) p=0.527		(108;115;104)
≤ 2 spinal fractures	-0.25 (-2.49, 2.00) p=0.829	0.975	1.39 (-1.17, 3.95) p=0.288	0.275	(77;95;86)
≥ 2 spinal fractures	-3.45 (-7.65, 0.74) p=0.111		0.13 (-4.75, 5.01) p=0.959		(52;49;49)
TLS at 12 months					
Female	3.94 (d-7.63,15.52) p=0.505	0.892	4.50 (-5.64, 14.64) p=0.384	0.112	(130,130,141)
Male	-5.91 (-51.19, 39.37) p=0.801		33.77 (-7.27, 74.82) p=0.119		(18;23;16)
Age ≤ 70	10.91 (-10.33,32.14) p=0.317	0.299	8.62 (-10.30,27.53) p=0.374	0.294	(58;59;64)
Age ≥ 70	2.07 (-10.99,15.14) p=0.756		8.34 (-3.36, 20.04) p=0.164		(90;94;93)
≤ 2 spinal fractures	8.17 (-9.44, 25.79) p=0.365	0.254	4.41 (-9.43,18.25) p=0.533	0.563	(76;93;81)
≥ 2 spinal fractures	-1.10 (-15.95,13.75) p=0.885		11.41 (-6.19, 29.02) p=0.207		(62;53;66)
QUALEFFO-41 at 4 months					
Female	-1.43 (-3.15, 0.29) p=0.103	0.531	-0.18 (-1.79, 1.43) p=0.827	0.138	(159,158,156)
Male	0.15 (-4.65, 4.96) p=0.950		4.08 (-3.23, 11.40) p=0.283		(21;27;17)
Age ≤ 70	-2.01 (-4.67, 0.65) p=0.141	0.134	-2.12 (-4.89, 0.64) p=0.135	0.204	(69;70;67)
Age ≥ 70	-0.91 (-3.06, 1.24) p=0.409		1.32 (-0.71, 3.34) p=0.204		(111;115;106)
≤ 2 spinal fractures	-1.04 (-2.99, 0.92) p=0.301	0.376	-0.76 (-3.00, 1.47) p=0.505	0.595	(82;103;88)

≥ 2 spinal fractures	-1.38 (-4.41, 1.64) p=0.373		1.39 (-1.76, 4.55) p=0.389		(62;58;53)
TLS at 4 months					
Female	7.05 (-2.52,16.61) p=0.150	0.702	9.29 (0.63,17.96) p=0.036	0.345	(144;147;144)
Male	7.43 (-17.87,32.74) p=0.570		8.67 (-17.04,34.38) p=0.514		(18;24;17)
Age ≤ 70	20.17 (5.89,34.4) p=0.007	0.008	21.09 (7.28,34.89) p=0.003	0.0006	(65;62;64)
Age ≥ 70	0.15 (-11.45,11,75) p=0.980		2.28 (-8.26,12.82) p=0.672		(97;109;97)
≤ 2 spinal fractures	9.81 (-4.48,24.09) p=0.18	0.195	6.97 (-4.78,18.73) p=0.247	0.294	(78;99;84)
≥ 2 spinal fractures	7.41 (-4.31,19.12) p=0.217		12.83 (-0.35,26.02) p=0.059		(76;67;71)

Table 4. Cost-effectiveness results using NHS and social services perspective at 12 months

		Complete case analysis (n=242)	Multiple imputation of costs and EQ5D* (n=613)
Exercise vs SSPT			
Difference in costs (£)	Mean	536	216
	SE	367	224
	95% CI	-183 to 1255	-223 to 655
Difference in QALYs adjusted for baseline EQ-5D	Mean	0.017	0.002
	SE	0.015	0.011
	95%CI	-0.013 to 0.048	-0.020 to 0.025
ICER	£/QALY	31,098	100,525
Probability of being CE at £20,000 per QALY		36%	31%
Manual vs SSPT			
Difference in costs (£)	Mean	-72	244
	SE	361	224
	95% CI	-780 to 635	-195 to 683
Difference in QALYs adjusted for baseline EQ-5D	Mean	-0.011	-0.014
	SE	0.015	0.012
	95%CI	-0.041 to 0.019	-0.036 to 0.009
ICER	£/QALY	6,710 (not cost-effective)	Dominated
Probability of being CE at £20,000 per QALY		39%	7%

*Multiple imputation by intervention with predictive mean matching using baseline covariates: female, age at randomisation, EQ5D score at baseline, mobility (walking distance), recruitment site, falls in previous year, back pain at baseline.

Figure 1: PROVE Study Consolidated Standards of Reporting Trials flow diagram